

Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Ivan Veinhardt Latták

Název práce Odvozování schématu v NoSQL databázích

Rok odevzdání 2021

Studijní program Informatika

Studijní obor Softwarové a datové inženýrství

Autor posudku Pavel Koupil

Role Vedoucí

Pracoviště Katedra softwarového inženýrství, MFF UK

Text posudku

Cílem diplomové práce bylo studium a vzájemné porovnání vybraných algoritmů odvozování schématu v dokumentových NoSQL databázových systémech. Na základě identifikace silných a slabých stránek těchto algoritmů měl být navržen nový algoritmus odvozování schématu nad dokumentovým modelem, který navíc bude rozšiřitelný pro práci nad dalšími datovými modely (umožňující schema-less přístup), a to zejména nad grafovými daty.

V první části práce se autor zabývá rešerší a porovnáním vybraných algoritmů na základě jejich statických vlastností. Součástí této části je také autorem velmi dobře navržená minimalistická kolekce JSON dokumentů, která je následně základem příkladů ilustrujících schémata odvozená jednotlivými algoritmy.

V druhé části autor vyčerpávajícím způsobem popisuje návrh vlastního přístupu. Autor se inspiroval existujícími přístupy, které ve své práci rozšiřuje a opravuje jejich nedostatky. Oproti vybraným přístupům navíc autor zvažuje kontext multi-modelových dat, a tedy možnou rozšiřitelnost nového přístupu a jeho aplikaci nad jinými datovými modely.

Následně autor provede experimentální porovnání vybraných algoritmů, a to 1) ověřením úplnosti odvozeného schématu a 2) porovnáním rychlostí odvození schématu. V rámci porovnání je vynechán jeden z algoritmů, neboť se včas nepodařilo získat zdrojové kódy tohoto přístupu, což ovšem není dáno chybou autora. Bohužel, výkon navrženého algoritmu v rámci provedených experimentů není nejlepší, ale s narůstající velikostí vstupních dat se výkonnostně přibližuje.

Autor se práci věnoval průběžně. Konzultace probíhaly pravidelně. Předložená práce splňuje požadavky kladené na diplomovou práci, a proto ji doporučuji k obhajobě. V případě, že autor ještě upraví výkon navrženého algoritmu a do experimentálního porovnání přidá chybějící přístup, práce má publikační potenciál.

Práci doporučuji k obhajobě.

Práci nenavrhují na zvláštní ocenění.

Datum 18. srpna 2021

Podpis